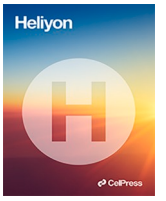
Heliyon 10 (2024) e29007

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Identification of qualitative characteristics of immunosuppression in sepsis based on immune-related genes and immune infiltration features

Ni Zeng, Zaijin Jian, Junmei Xu, Tian Peng, Guiping Hong, Feng Xiao *

Department of Anesthesiology, The Second Xiangya Hospital, Central South University, Changsha, 410011, China

A B S T R A C T

Objective: Sepsis is linked to high morbidity and mortality rates. Consequently, early diagnosis is crucial for proper treatment, reducing hospitalization, and mortality rates. Additionally, over one-fifth of sepsis patients still face a risk of death. Hence, early diagnosis, and effective treatment play pivotal roles in enhancing the prognosis of patients with sepsis.

Method: The study analyzed whole blood data obtained from patients with sepsis and control samples sourced from three datasets (GSE57065, GSE69528, and GSE28750). Commonly dysregulated immune-related genes (IRGs) among these three datasets were identified. The differential characteristics of these common IRGs in the sepsis and control samples were assessed using the REO-based algorithm. Based on these differential characteristics, samples from eight Gene Expression Omnibus (GEO) databases (GSE57065, GSE69528, GSE28750, GSE65682, GSE69063, GSE95233, GSE131761, and GSE154918), along with three ArrayExpress databases (E-MTAB-4421, E-MTAB-4451, and E-MTAB-7581), were categorized and scored. The effectiveness of these differential characteristics in distinguishing sepsis samples from control samples was evaluated using the AUC value derived from the receiver operating characteristic curve (ROC) curve. Furthermore, the expression of IRGs was validated in peripheral blood samples obtained from patients with sepsis through qRT-PCR.

Results: Among the three training datasets, a total of 84 common dysregulated immune-related genes (IRGs) were identified. Utilizing a within-sample relative expression ordering (REOs)-based algorithm to analyze these common IRGs, differential characteristics were observed in three reverse stable pairs (ELANE-RORA, IL18RAP-CD247, and IL1R1-CD28). In the eight GEO datasets, the expression of ELANE, IL18RAP, and IL1R1 demonstrated significant upregulation, while RORA, CD247, and CD28 expression exhibited notable downregulation during sepsis. These three pairs of immune-related marker genes displayed accuracies of 95.89% and 97.99% in distinguishing sepsis samples among the eight GEO datasets and the three independent ArrayExpress datasets, respectively. The area under the receiver operating characteristic curve ranged from 0.81 to 1.0. Additionally, among these three immune-related marker gene pairs, mRNA expression levels of ELANE and IL1R1 were upregulated, whereas the levels of CD247 and CD28 mRNA were downregulated in blood samples from patients with sepsis compared to normal controls.

Conclusion: These three immune-related marker gene pairs exhibit high predictive performance for blood samples from patients with sepsis. They hold potential as valuable auxiliary clinical blood screening tools for sepsis.

1. Introduction

Sepsis, a common complication following traumatic injury, manifests as life-threatening organ dysfunction due to the dysregulated host response to infection [1]. Developed countries witness an incidence rate of around 0.1% (100 sepsis cases per 100,000 individuals), with approximately 2% of hospitalized patients developing sepsis upon admission [1,2]. Despite significant advancements in sepsis management, more than one-fifth of sepsis cases remain at risk of mortality [3]. Hence, early diagnosis, and effective

* Corresponding author.
E-mail address: xiaofengcsu@csu.edu.cn (F. Xiao).

https://doi.org/10.1016/j.heliyon.2024.e29007
Received 31 May 2023; Received in revised form 26 March 2024; Accepted 28 March 2024
Available online 3 April 2024

treatment are crucial, given that every hour delay in sepsis treatment heightens the risk of death [4]. Research has shown that biomarkers can predict disease activity and progression in the early stages, guiding treatment planning and evaluation of treatment response [5]. However, there is a notable absence of established biomarkers for diagnosing, prognosing, and detecting sepsis [6]. Therefore, it is imperative and of utmost importance to identify biomarkers capable of facilitating the prediction and early diagnosis of sepsis.

Several studies have proposed that various risk factors, such as thrombocytopenia [7,8] and alcohol abuse [9], contribute to the mortality of sepsis patients. Immunosuppression, particularly post-clinical 'recovery,' has been identified as a crucial factor [10–13]. Previous evidence has demonstrated that analyzing the differential expression of immune-related genes (IRGs) in sepsis aids in identifying biomarkers for early prediction and diagnosis [14–18]. Therefore, IRGs hold promise as potential biomarkers for predicting and treating sepsis. These genes are intricately linked to the immune response, and their expression patterns can serve as valuable indicators for predicting sepsis, devising treatment strategies, and monitoring treatment response [14,19,20]. Recently, High-throughput genome sequencing technologies (HTS) have emerged, aimed at identifying differentially expressed genes (DEGs) between two phenotypes [21]. A prior study highlighted HTS application in detecting IRG signatures in cancer [22]. Furthermore, differentially expressed IRGs (DEIRGs) in sepsis can be identified using HTS, offering potential biomarkers for early sepsis diagnosis [14]. However, because of an insufficient number of samples, it is difficult to identify DEGs with slight expression differences in practical situations [23].

While combining multiple datasets from various laboratories aids in identifying DEGs with slight expression differences, direct integration of these datasets might encounter interference from various random factors (such as batch effects) [24]. In addition to batch effects, several other random factors complicate this process, including experimental and sampling variability, measurement noise, data processing variability, and subject variability. The inherent variability in biological samples arises from the stochastic nature of biological processes. Random fluctuations in gene expression, cell populations, and molecular interactions contribute to this variability, making it challenging to discern genuine biological changes from natural fluctuations [25]. Furthermore, inconsistencies in data preprocessing steps, like data normalization and transformation, can introduce additional variability if not uniformly applied across datasets. In datasets containing clinical data, individual subject variability acts as another random factor, where each patient may respond uniquely to treatment or exhibit distinct physiological profiles, adding randomness to the data [26]. These factors limit the clinical utility of quantitative transcriptional characterization. Nevertheless, studies have reported that the within-sample relative expression orderings (REOs) of gene pairs exhibit robustness against experimental batch effects [23,27]. Additionally, accumulating evidence showcases successful applications of REOs in diagnosing various diseases, such as cancer [28] and Alzheimer's disease [29]. Hence, leveraging within-sample REOs of genes may offer a means to identify robust qualitative characteristics of immunosuppression for early sepsis diagnosis.

By focusing on immune-related genes, our approach holds broader applicability across diverse patient demographics, considering the universal role of the immune response in sepsis. For instance, ELANE's involvement in the selective killing of neutrophils by multiple human cancer cells highlights its significance [30]. Similarly, RORA expression positively correlates with immune cell infiltration, particularly notable in lung adenocarcinoma (LUAD) [31]. Aberrant expression of IL18RAP has been observed across various cancers, influencing tumor immunity and yielding diverse clinical outcomes [32]. CD247, implicated in multiple autoimmune diseases, plays a pivotal role in a shared mechanism mediated by T cells [33]. In pancreatic ductal adenocarcinoma, IL1R1's association with location and immune response is noteworthy; elevated expression in pancreatic head cancer correlates with higher immunological scores and predicts an unfavorable outcome [34]. Moreover, the role of CD28 in the survival of regulatory T cells and maintenance of immunological homeostasis is of paramount importance [35]. Our study aims to develop a diagnostic model for sepsis, intending to function as an auxiliary clinical blood screening tool.

2. Material and methods

2.1. Data processing

Eight datasets [three training datasets (GSE28750, GSE69528, and GSE57065] and five validation datasets (GSE65682, GSE69063,

Table 1
Datasets used in this study.

Datasets	Control Sample	Mortality/Sepsis sample	Type
GSE28750	20	−/10	whole blood
GSE69528	55	−/83	whole blood
GSE57065	25	−/82	whole blood
GSE65682	42	48/231	whole blood
GSE69063	33	−/57	peripheral blood
GSE95233	22	34/102	whole blood
GSE131761	15	−/114	whole blood
GSE154918	40	−/53	whole blood
E-MTAB-4421	–	58/270	leukocyte
E-MTAB-4451	–	57/114	leukocyte
E-MTAB-7581	–	48/176	whole blood

GSE95233, GSE131761, and GSE154918]) were downloaded from Gene Expression Omnibus (GEO) (https://www.ncbi.nlm.nih.gov/geo/). Three datasets (E-MTAB-4421, E-MTAB-4451, and E-MTAB-7581) were downloaded from ArrayExpress (https://www. ebi. ac. uk/ArrayExpress). The datasets used in this study are listed in Table 1. ImmPort (The Immunology Database and Analysis Portal database, http://www.immport.org/) was used to select immune-related gene (IRGs) sets in immunology research, containing 2489 IRGs. DEIRGs were analyzed using the limma package in R. All data were analyzed using Python (v3.7.6) and R Studio (v4.0.2).

2.2. REOs-based algorithm

Stable gene pairs obtained from the training datasets can serve as classification markers for any individual sample. A classification signature was obtained based on stable gene pairs. For unclassified samples, the REOs of all stable gene pairs were calculated. The comparison of two genes in a gene pair ($_A$, $_B$) was viewed as an event with only two possible outcomes: the expression value of gene A (G_A) was either higher or lower than that of gene B (G_B), and REO was denoted as $G_A > G_B$ or $G_A < G_B$. The probability of occurrence of $G_A < G_B$ in n stable gene pairs of a sample fewer than or equal to k times was calculated using the following formula:

Binomial distribution:

$$p = 1 - \sum_{i=0}^{k-1} \binom{n}{i} 0.5^i (1 - 0.5)^{n-i}$$

where n is the number of samples and k is the number of occurrences of $G_A < G_B$ in n samples. A p-value of less than 0.05 indicates that $G_A < G_B$ can be maintained in most samples, and it's defined as highly stable pairs.

The reversal pattern of stable gene pairs can also be used to classify two types of samples (categories I and II). If the gene pair (G_A, G_B) is stable in both groups, but $G_A < G_B$ in one group and $G_A > G_B$ in another, the gene pair (G_A, G_B) is called a stable reversal pair. This method is also applicable to binomial distributions to test the probability of samples belonging to each category.

For a sample with an unknown category, k is set as the number of times $G_A > G_B$ appears in a stable reversal pair, and n is the number of highly stable pairs. where k denotes the number of stable pairs. If $P(k, n) < 0.05$, the sample notably conformed to the characteristics of Category I and was classified as Category I. If $P(n\text{-}k, n) < 0.05$, the sample conforms to the characteristics of Category II, and is classified as Category II.

2.3. Exploring optimal score thresholds for discriminating sepsis samples from control samples

The expression patterns of the three gene pairs were utilized to assign scores to all samples. The presence of $G_A < G_B$ indicated a score of 1 [36], signifying a bias towards sepsis group characteristics. Conversely, samples lacking the reversal gene pair received a score of −1. Based on the overall score bias, samples were classified into either the normal or sepsis groups. Rigorous concordance analysis was conducted on both the training and validation sets to determine optimal score thresholds, ensuring the highest levels of accuracy.

2.4. Blood sample collection

Blood samples were obtained from patients with sepsis (n = 13) and from healthy donors (n = 15). The modified early warning sign-sepsis recognition system (MEWS-SRS) was used to screen for sepsis [37]. Patients enrolled in the study received standardized care for sepsis following the guidelines established by the Surviving Sepsis Consensus. Informed consent was obtained from all participants contributing blood samples. The collection of blood samples was conducted with approval from the Second Xiangya Hospital (No.2023[Clinical Research]091).

2.5. qRT-PCR

Total RNA was extracted using TRIzol reagent (Thermo Fisher Scientific), and a quantity of 1000 ng RNA was used as a template for the synthesis of first-strand cDNA. Amplification was performed using a SYBR Green Supermix kit (Bio-Rad, Hercules, CA, USA) [38]. The relative expression of ELANE, IL18RAP, IL1R1, RORA, CD247, and CD28 was determined using the $2^{-\Delta\Delta CT}$ technique, with β-tectin serving as the internal reference. Primers used are listed in Supplementary Table S1.

2.6. Statistical analysis

Statistical analyses were conducted using Python (v3.7.6) and R Studio (v4.0.2). The binary classification performance of these characteristics was assessed using the area under the receiver operating characteristic curve (ROC) curve (AUC). For comparisons between two groups, a t-test was employed for statistical analysis. A p-value of less than 0.05 indicated a statistically significant difference.

3. Results

3.1. Sepsis signatures based on DEIRGs were constructed using the REOs-based algorithm

The analysis process is illustrated in Fig. 1. First, DEIRGs from three datasets (GS E57065, GSE69528, and GSE28750) were identified (Fig. 2A–F). A total of 84 common DEIRGs were identified after comparing the results of the three datasets (Fig. 2G). Furthermore, the differential characteristics between sepsis and control samples were analyzed using the REO-based algorithm; 35, 28, and 27 reverse gene pairs were obtained from GSE57065-GSE69528, GSE57065-GSE28750, and GSE69528-GSE28750, respectively (Table 2). Among them, there were 3 reverse gene pairs (ELANE-RORA, IL18RAP-CD247, and IL1R1-CD28) overlapped in GSE57065-GSE69528, GSE57065-GSE28750, and GSE69528-GSE28750 (Fig. 2H) (Table 3). The expression levels of the six genes in the three datasets are listed in Table 4.

3.2. Expression characteristics of three immune-related marker gene pairs in sepsis and control samples were analyzed

Subsequently, the expression of these 3 immune-related marker gene pairs (ELANE-RORA, IL18RAP-CD247, and IL1R1-CD28) for sepsis in three training datasets (GSE57065, GSE69528, and GSE28750) was analyzed. According to the results, ELANE, IL18RAP, and IL1R1 expressions were notably upregulated, while RORA, CD247, and CD28 expressions were remarkably downregulated in sepsis patients (Fig. 3A). Furthermore, the expression of these three sepsis immune-related marker gene pairs was validated in the other five datasets (GSE65682, GSE69063, GSE95233, GSE131761, and GSE154918), and consistent findings were observed in the five validation datasets (Fig. 3B).

3.3. The performance of three immune-related marker gene pairs in classifying sepsis and control samples was evaluated

Furthermore, based on the above three immune-related marker gene pairs, sepsis-associated datasets were scored. As shown by the results, in the three training datasets (GSE57065, GSE69528, and GSE28750), when the sample score threshold was higher than 2, the sepsis and control samples could be noticeably distinguished (Fig. 4A). Moreover, in the other five independent validation datasets containing sepsis and normal samples, when the score threshold was higher than two, the immune-related marker gene pairs showed the best classification performance for sepsis and normal samples (Fig. 4B). Based on these findings, the classification accuracy of immune-related marker gene pairs was analyzed. These immune-related marker gene pairs had an accuracy of 95.89% in distinguishing sepsis samples from eight GEO datasets (1240 sepsis samples in total) (Fig. 4C and D) (Table 5). Additionally, these marker gene pairs showed an accuracy of 97.99% in distinguishing sepsis samples from three independent ArrayExpress datasets (547 sepsis

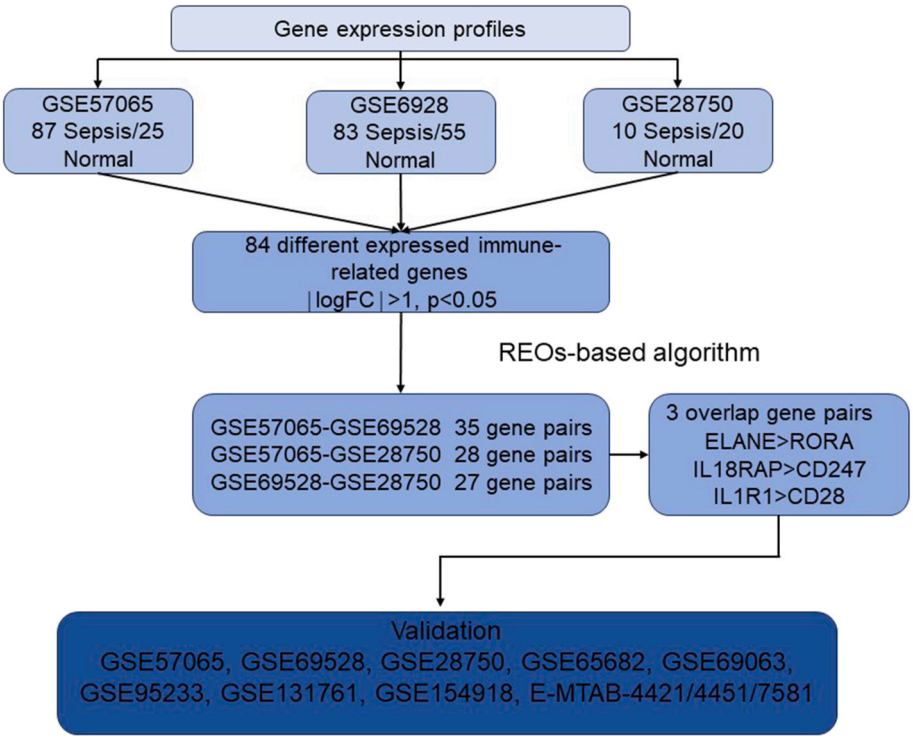


Fig. 1. Flow chart showing the analysis process.

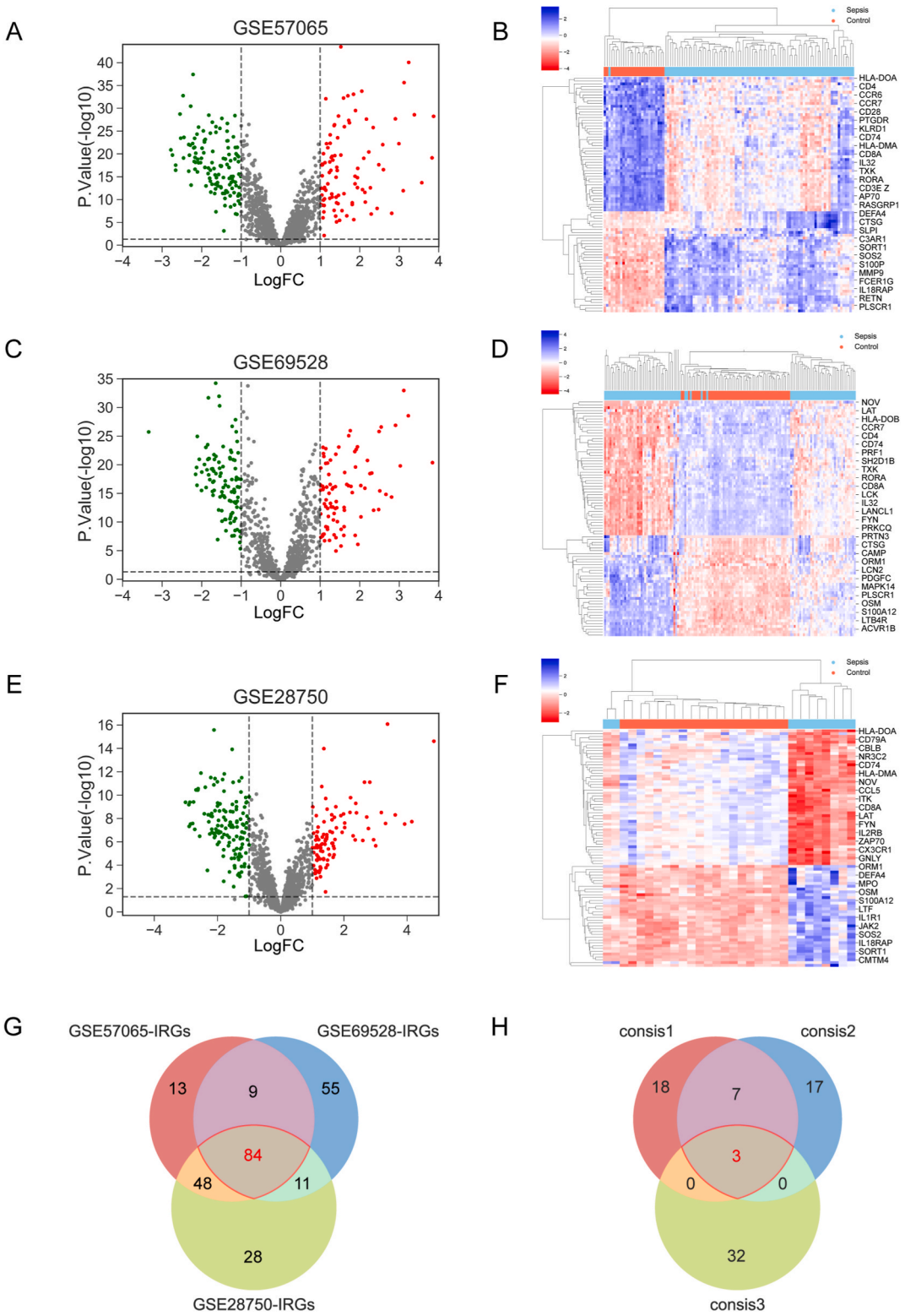


Fig. 2. Sepsis signature based on DEIRGs was constructed using the REOs-based algorithm. (A–F) The DEIRGs in the GSE57065, GSE69528, and GSE28750 datasets were shown using a volcano map and heat map, respectively. The cut-off value is |logFC|>1, $p < 0.05$. (G) The common DEIRGs in the GSE57065, GSE69528, and GSE28750 datasets were shown using a Venn diagram. (H) The overlapped reverse gene pairs among GSE57065-GSE69528, GSE57065-GSE28750, and GSE69528-GSE28750 were screened using a Venn diagram.

Table 2
Top50 DEIRGs with REOs.

Dataset1	Dataset2	Pair_up	Pair_down	Reverse	Consis	Overlap
GSE57065	GSE69528	48	40	296	35	3
GSE57065	GSE28750	49	48	193	28	
GSE69528	GSE28750	48	41	121	27	

Table 3
The REO-based signature.

Gene pair	REO (Sepsis vs Normal)
pair1	ELANE > RORA
pair2	IL18RAP > CD247
pair3	IL1R1>CD28

Table 4
The expression levels of ELANE, RORA, IL18RAP, CD247, IL1R1 and CD28 in GSE57065, GSE69528, and GSE28750.

gene	data	logFC	reg	P	adjP
CD247	GSE57065	−2.55	Down	1.76E-29	9.74E-27
IL18RAP	GSE57065	2.17	Up	2.34E-28	9.06E-26
IL1R1	GSE57065	2.23	Up	3.61E-21	2.4E-19
RORA	GSE57065	−1.13	Down	9.51E-21	5.8E-19
CD28	GSE57065	−1.84	Down	2E-16	5.61E-15
ELANE	GSE57065	2.44	Up	8.67E-09	6.98E-08
CD247	GSE69528	−1.82486014	Down	8.55E-22	2.09E-20
CD28	GSE69528	−1.20224808	Down	1.31E-21	3.09E-20
RORA	GSE69528	−1.86619526	Down	7.54E-20	1.3E-18
IL18RAP	GSE69528	1.71408866	Up	4.46E-17	4.84E-16
IL1R1	GSE69528	1.40552635	Up	2.73E-14	1.98E-13
ELANE	GSE69528	2.35119625	Up	8.96E-13	5.3E-12
CD247	GSE28750	−2.68178561	Down	3.52E-11	1.32E-08
IL18RAP	GSE28750	1.98816577	Up	3.99E-09	0.00000033
RORA	GSE28750	−1.7462097	Down	1.31E-08	0.00000078
IL1R1	GSE28750	2.02334643	Up	2.43E-08	0.00000124
CD28	GSE28750	−1.9889238	Down	3.97E-08	0.00000177
ELANE	GSE28750	2.35498781	Up	0.000000673	0.0000164

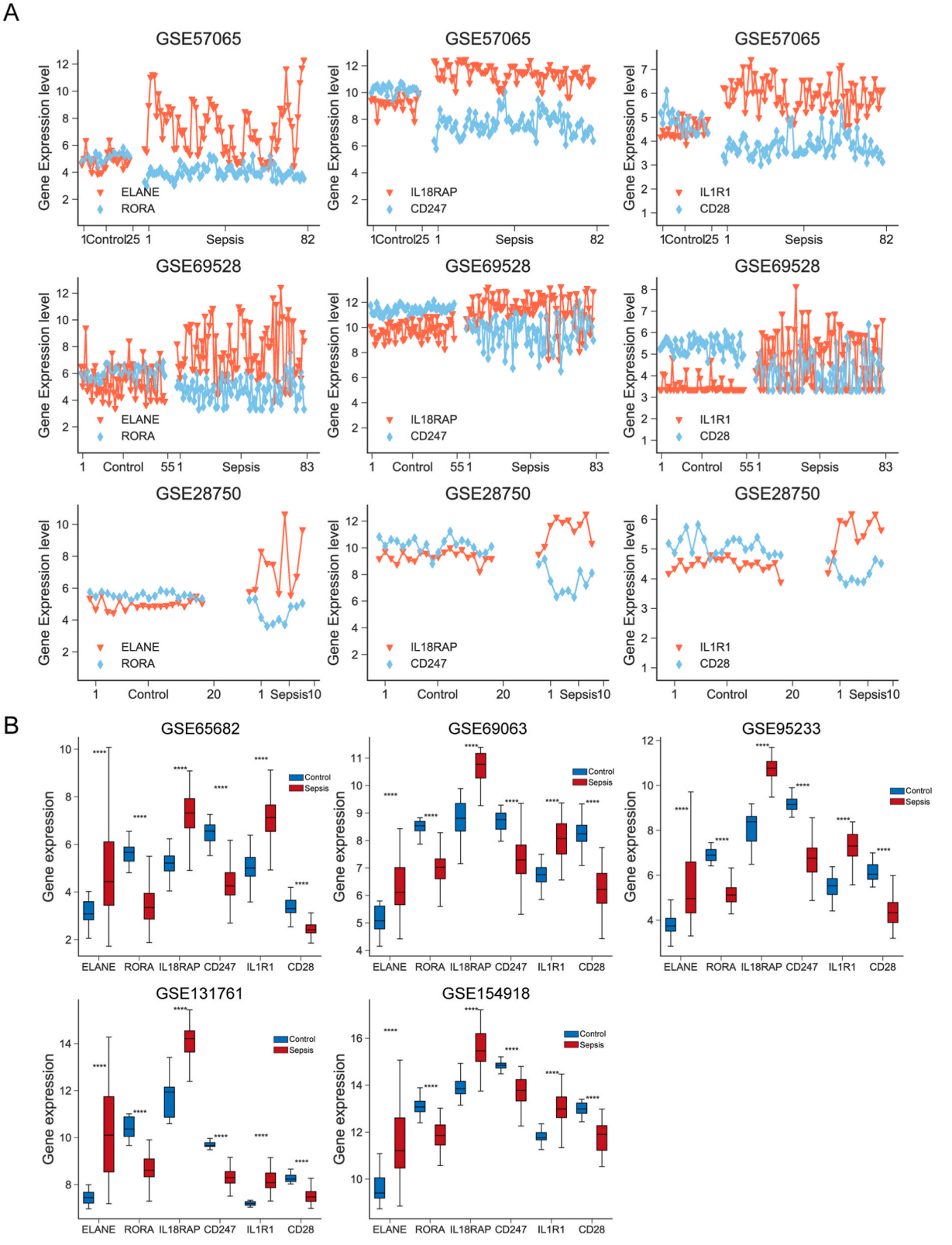
samples in total) (Fig. 4D) (Table 5). Finally, qualitative diagnostic characteristics were validated in 8 GEO datasets through ROC and AUC analysis. According to the results, the AUC values ranged from 0.82 to 1.00, which indicated that these three immune-related marker gene pairs could effectively classify patients with sepsis and normal controls (Fig. 4E). The ROC curves for each gene are shownin Fig. S1. Taken together, the differential characteristics (composed of three immune-related marker gene pairs) showed a high predictive performance for distinguishing the blood samples of patients with sepsis from those of normal controls.

3.4. Expression of the six IRGs in sepsis patients' peripheral blood samples

To further explore the roles of these IRGs, blood samples were collected from both patients with sepsis and healthy controls, and the mRNA expression levels of ELANE, IL18RAP, IL1R1, RORA, CD247, and CD28 were measured. As depicted in Fig. 5, the mRNA expression levels of ELANE and IL1R1 were notably upregulated, whereas the mRNA expression levels of CD247 and CD28 were significantly downregulated in the blood samples of patients with sepsis when compared to those of the normal controls (Fig.5A and D-F). Through RORA mRNA level lower and IL18RAP mRNA level higher in the sepsis group, there is no statistical difference (Fig. 5B and C).

4. Discussion

In this study, three immune-related marker gene pairs efficiently distinguished septic samples from control samples with high accuracy, and their potential in clinically assisting the screening of early sepsis was validated. After analyzing the three GEO datasets, 84 DEIRGs were identified. Furthermore, the REO-based algorithm was used to analyze the GEO dataset pairs (GSE57065-GSE69528, GSE57065-GSE28750, and GSE69528-GSE28750), and the differential characteristics of three reverse gene pairs (ELANE-RORA, IL18RAP-CD247, and IL1R1-CD28) were identified. Several studies have shown that these genes are closely associated with sepsis. For example, ELANE has been identified as a component of the diagnostic gene signatures in several studies [39–41]. In addition, ELANE serves as a target of miR-608 in human monocytes and participates in the process of inflammation; therefore, it is a novel target for the



(caption on next page)

Fig. 3. Expression characteristics of three immune-related marker gene pairs in sepsis and control samples were analyzed. (A) The expressions of 3 immune-related marker gene pairs (ELANE-RORA, IL18RAP-CD247, and IL1R1-CD28) in datasets GSE57065, GSE69528, and GSE28750) were analyzed. (B) The expressions of 3 immune-related marker gene pairs for sepsis were validated in the other 5 datasets (GSE65682, GSE69063, GSE95233, GSE131761, and GSE154918).

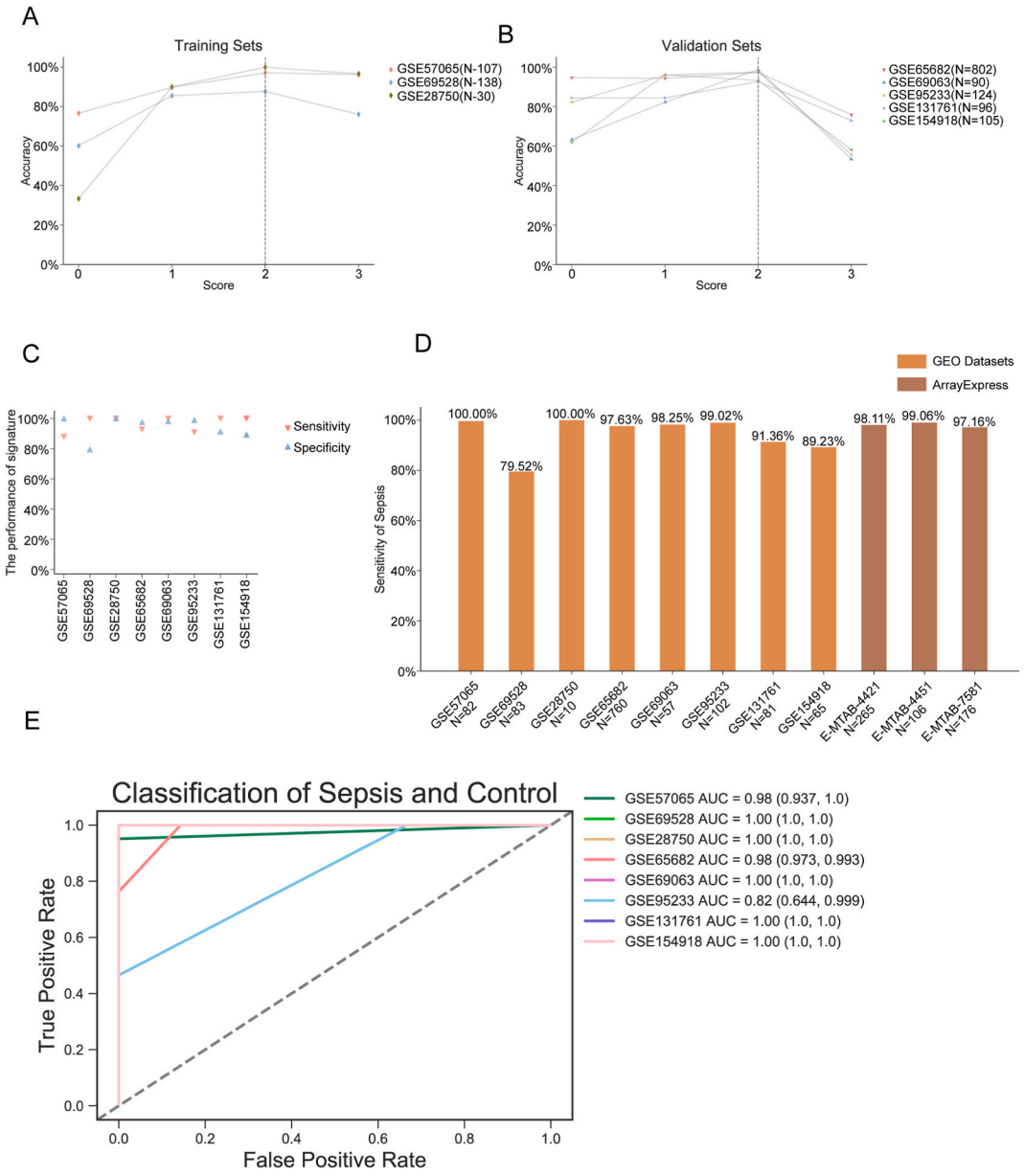


Fig. 4. The performance of three immune-related marker gene pairs in classifying sepsis and control samples was evaluated. (A) The performance of 3 immune-related marker gene pairs for distinguishing sepsis samples from control samples in 3 training datasets GSE57065, GSE69528, and GSE28750 was evaluated. (B) The performance of 3 immune-related marker gene pairs for distinguishing sepsis samples from control samples in 5 validation datasets GSE65682, GSE69063, GSE95233, GSE131761, and GSE154918 was evaluated. (C) The sensitivity and specificity of marker gene pairs in 8 GEO datasets (3 training datasets and 5 validation datasets) were analyzed. (D) The sensitivity and specificity of marker gene pairs in 8 GEO datasets and 3 ArrayExpress datasets were analyzed. (E) The qualitative diagnostic characteristics were validated in 8 GEO datasets through ROC and AUC analysis.

diagnosis or treatment of sepsis [42]. RORA is implicated in the context of the inflammatory response and can drive pathological transcriptional states, with implications for sepsis biology and treatment [43]. The regulation of RORA expression plays a role in the treatment of septic response in mice [44]; selective inhibition of RORA helps reduce the severity of lipopolysaccharide (LPS)-induced

Table 5
The classification performance of 3 immune-related marker gene pairs for sepsis samples and control samples with score ≥2.

Source	Cohort	Sepsis sample	Score≥2	Positive rate
GEO datasets	GSE57065	82	82	100.00%
	GSE69528	83	66	79.52%
	GSE28750	10	10	100.00%
	GSE65682	760	742	97.63%
	GSE69063	57	56	98.25%
	GSE95233	102	101	99.02%
	GSE131761	81	74	91.36%
	GSE154918	65	58	89.23%
	Total	1240	1189	95.89%
ArrayExpress	E-MTAB-4421	265	260	98.11%
	E-MTAB-4451	106	105	99.06%
	E-MTAB-7581	176	171	97.16%
	Total	547	536	97.99%
Total		1787	1725	96.53%

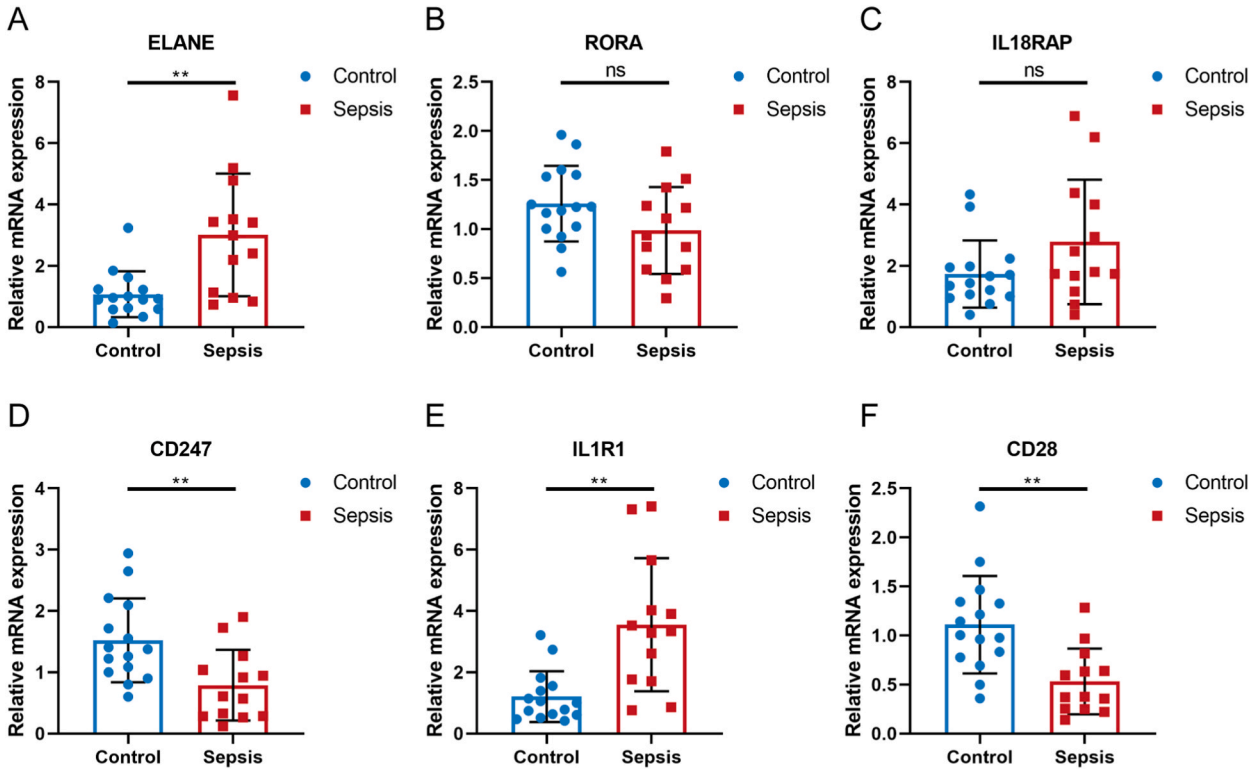


Fig. 5. Expression of the six factors in sepsis patients (A–F) Blood samples were collected from sepsis patients (n = 13) and normal control (n = 15) and examined for the mRNA expression levels of ELANE, IL18RAP, IL1R1, RORA, CD247, and CD28 using qRT-PCR.

septic shock [45]. IL18RAP is closely associated with immune-related biological processes and sepsis [46]. CD247 is poorly expressed in septic mice [47] and is a promising target for the prognosis of sepsis in LPS-induced septic mice [48]. IL1R1 is upregulated in microglial cells of rats with sepsis-associated encephalopathy [49]; inhibiting IL1R1 can attenuate liver injury in septic mice [50]. CD28 expression is abnormally downregulated in sepsis, and activation of CD28 helps to improve the survival rate of septic mice [51]. In this study, the expression of differential characteristics (composed of three immune-related marker gene pairs) in three GEO training datasets and five GEO validation datasets was analyzed. Consistent results have been reported in previous studies. Thus, the differential characteristics screened using the REO-based algorithm may serve as the qualitative characteristics of sepsis.

After three immune-related gene pairs, their performance in distinguishing sepsis samples from control samples was evaluated. It was found that the differential characteristics (composed of three sepsis immune-related gene pairs) could efficiently classify sepsis and control samples in eight GEO datasets and three ArrayExpress datasets with accuracies of 95.89% and 97.99%. The AUC value ranged from 0.8 to 1.0, which indicated its high classification efficiency. Therefore, these differential characteristics (composed of 3 immune-related gene pairs) are potential qualitative characteristics of sepsis. A previous study has provided evidence for the use of these

diagnostic biomarkers for sepsis. Yao et al. reported a four-gene signature (ANXA3, CD177, GRAMD1C, and TIGD3) to distinguish patients with pediatric sepsis based on the GSE119217 dataset [52]. Through analysis of three sepsis-associated datasets (GSE95233, GSE57065, and GSE28750), Gong et al. have identified 9 genes (LRG1, ELANE, TP53, LCK, TBX21, ZAP70, CD247, ITK, and FYN) as potential new biomarkers for sepsis [53]. Although several studies have proposed potential genetic markers for sepsis diagnosis, their sample sizes are relatively small, and their results have not been sufficiently verified on other datasets. However, in the present study, 1787 samples from different data platforms (composed of GSE datasets and ArrayExpress) were validated. Therefore, this study overcomes the shortcomings of previous studies. Nevertheless, the 6 IRGs have only been validated in a small clinical sample size. Validation with a larger sample size is required in the future.

In summary, this study, for the first time, utilized an REO-based algorithm to predict differential characteristics for distinguishing sepsis from control samples. These differential characteristics (composed of 3 sepsis immune-related marker gene pairs) exhibited robust classification accuracy in different sepsis-related datasets. These findings indicate that REO-based algorithm-predicted differential characteristics can be unaffected by batch effects, which may serve as an assistant tool for the early screening of sepsis in the clinic. In the domain of sepsis diagnosis using immune-related genes, the significant advantages include a high level of accuracy in discerning sepsis samples and the prospective use of tailored treatment. Nevertheless, obstacles remain when modern bioinformatics methods are used, such as the intrinsic diversity of sepsis and the potential for overfitting. It is important to acknowledge and tackle these problems to enhance the applicability and dependability of the results in this field of study. Compared to previous studies, our work introduces a novel approach for the early diagnosis of sepsis by employing an REO-based algorithm. Lu et al. [14] utilized modified Lasso penalized regression and Random Forest to construct an IRG classifier, and Wang et al. [54] focused on pediatric sepsis using various bioinformatics analyses; our study is the first to identify three immune-related gene pairs using an REO-based algorithm. Additionally, unlike Peng et al. [20], whose focus was on predicting 28-day mortality in patients with sepsis, our study aimed at early diagnosis in a broader population with a substantially larger sample size. The robustness of our method against experimental batch effects significantly bolsters its potential for clinical applications in early sepsis screening. However, our study faces several notable limitations. The absence of prospective clinical validation using a large sample size limits the comprehensive establishment of the practical utility of the identified biomarkers. Moreover, the incorporation of retrospective gene expression datasets and the utilization of expression levels of six factors in clinical blood samples might introduce potential biases, potentially failing to fully capture the intricate and dynamic nature of sepsis cases, along with potential confounding variables.

5. Conclusion

Our innovative REO-based method for early sepsis detection has exhibited robustness, accurately distinguishing sepsis samples from controls. This approach holds considerable promise for widespread clinical application, potentially benefiting healthcare professionals by enabling precise sepsis screening and timely interventions. As we continue to refine and validate this method, its potential to significantly enhance sepsis diagnosis and treatment becomes increasingly apparent, ultimately improving patient prognosis.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and materials

The authors confirm that the data supporting the findings of this study are available within the article and its supplementary materials.

Funding

This work has been supported by the National Natural Science Foundation of China (No.81971819) and the Hunan Provincial Natural Science Foundation (No. 2020JJ4777).

CRediT authorship contribution statement

Ni Zeng: Investigation, Conceptualization, Writing – original draft. **Zaijin Jian:** Investigation, Software. **Junmei Xu:** Software, Investigation. **Tian Peng:** Investigation. **Guiping Hong:** Investigation. **Feng Xiao:** Writing – review & editing, Supervision, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to

influence the work reported in this paper.

Acknowledgement

None.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e29007.

References

[1] M. Singer, C.S. Deutschman, C.W. Seymour, M. Shankar-Hari, D. Annane, M. Bauer, et al., The third international consensus definitions for sepsis and septic shock (Sepsis-3), JAMA 315 (2016) 801–810.
[2] D.C. Angus, T. van der Poll, Severe sepsis and septic shock, N. Engl. J. Med. 369 (2013) 840–851.
[3] M.M. Levy, A. Artigas, G.S. Phillips, A. Rhodes, R. Beale, T. Osborn, et al., Outcomes of the Surviving Sepsis Campaign in intensive care units in the USA and Europe: a prospective cohort study, Lancet Infect. Dis. 12 (2012) 919–924.
[4] R. Ferrer, I. Martin-Loeches, G. Phillips, T.M. Osborn, S. Townsend, R.P. Dellinger, et al., Empiric antibiotic treatment reduces mortality in severe sepsis and septic shock from the first hour: results from a guideline-based performance improvement program, Crit. Care Med. 42 (2014) 1749–1755.
[5] A. Almangush, I. Heikkinen, A.A. Mäkitie, R.D. Coletta, E. Läärä, I. Leivo, et al., Prognostic biomarkers for oral tongue squamous cell carcinoma: a systematic review and meta-analysis, Br. J. Cancer 117 (2017) 856–866.
[6] J.U. Jensen, L. Bouadma, Why biomarkers failed in sepsis, Intensive Care Med. 42 (2016) 2049–2051.
[7] A. Assinger, W.C. Schrottmaier, M. Salzmann, J. Rayes, Platelets in sepsis: an update on experimental models and clinical data, Front. Immunol. 10 (2019) 1687.
[8] B. Burunsuzoglu, C. Salturk, Z. Karakurt, E.A. Ongel, H.B. Takir, F. Kargin, et al., Thrombocytopenia: a risk factor of mortality for patients with sepsis in the intensive care unit, Turk Thorac J 17 (2016) 7–14.
[9] A. Gandhirajan, S. Roychowdhury, C. Kibler, E. Cross, S. Abraham, A. Bellar, et al., SIRT2-PFKP interaction dysregulates phagocytosis in macrophages with acute ethanol-exposure, Front. Immunol. 13 (2022) 1079962.
[10] M.J. Delano, P.A. Ward, Sepsis-induced immune dysfunction: can immune therapies reduce mortality? J. Clin. Invest. 126 (2016) 23–31.
[11] D. Liu, S.Y. Huang, J.H. Sun, H.C. Zhang, Q.L. Cai, C. Gao, et al., Sepsis-induced immunosuppression: mechanisms, diagnosis and current treatment options, Mil Med Res 9 (2022) 56.
[12] K.M. Sundar, M. Sires, Sepsis induced immunosuppression: implications for secondary infections and complications, Indian J. Crit. Care Med. 17 (2013) 162–169.
[13] F. Venet, G. Monneret, Advances in the understanding and treatment of sepsis-induced immunosuppression, Nat. Rev. Nephrol. 14 (2018) 121–137.
[14] J. Lu, R. Chen, Y. Ou, Q. Jiang, L. Wang, G. Liu, et al., Characterization of immune-related genes andimmune infiltration features for early diagnosis, prognosis and recognition of immunosuppression in sepsis, Int Immunopharmacol 107 (2022) 108650.
[15] Y. Yang, Y. Zhang, S. Li, X. Zheng, M.H. Wong, K.S. Leung, et al., A robust and generalizable immune-related signature for sepsis diagnostics, IEEE ACM Trans. Comput. Biol. Bioinf 19 (2022) 3246–3254.
[16] J. Niu, B. Qin, C. Wang, C. Chen, J. Yang, H. Shao, Identification of key immune-related genes in the progression of septic shock, Front. Genet. 12 (2021) 668527.
[17] Z.H. Chen, W.Y. Zhang, H. Ye, Y.Q. Guo, K. Zhang, X.M. Fang, A signature of immune-related genes correlating with clinical prognosis and immune microenvironment in sepsis, BMC Bioinf. 24 (2023) 20.
[18] Y. Tian, C. Wang, Q. Lu, C. Zhang, L. Hu, J. Ling, et al., Screening of potential immune-related genes expressed during sepsis using gene sequencing technology, Sci. Rep. 13 (2023) 4258.
[19] Y. Zheng, B. Liu, X. Deng, Y. Chen, Y. Huang, Y. Zhang, et al., Construction and validation of a robust prognostic model based on immune features in sepsis, Front. Immunol. 13 (2022) 994295.
[20] Y. Peng, Q. Wu, H. Liu, J. Zhang, Q. Han, F. Yin, et al., An immune-related gene signature predicts the 28-day mortality in patients with sepsis, Front. Immunol. 14 (2023) 1152117.
[21] D. Ou, Y. Wu, The prognostic and clinical significance of IFI44L aberrant downregulation in patients with oral squamous cell carcinoma, BMC Cancer 21 (2021) 1327.
[22] B. Chen, C. Hu, L. Jiang, Z. Xiang, Z. Zuo, Y. Lin, et al., Exploring the significance of novel immune-related gene signatures in the prognosis and immune features of pancreatic adenocarcinoma, Int Immunopharmacol 92 (2021) 107359.
[23] H. Liu, Y. Li, J. He, Q. Guan, R. Chen, H. Yan, et al., Robust transcriptional signatures for low-input RNA samples based on relative expression orderings, BMC Genom. 18 (2017) 913.
[24] J.T. Leek, R.B. Scharpf, H.C. Bravo, D. Simcha, B. Langmead, W.E. Johnson, et al., Tackling the widespread and critical impact of batch effects in high-throughput data, Nat. Rev. Genet. 11 (2010) 733–739.
[25] A. Swain, L.E. Azevedo-Schmidt, S.A. Maccracken, E.D. Currano, J.A. Dunne, C.C. Labandeira, et al., Sampling bias and the robustness of ecological metrics for plant-damage-type association networks, Ecology 104 (2023) e3922.
[26] J.F. Tulloch, Bias and variability in clinical research, Clin. Orthod. Res. 1 (1998) 94–96.
[27] X. Chen, X. Guo, P. He, J. Nie, X. Yan, J. Zhu, et al., Interactive influence of N6AMT1 and As3MT genetic variations on arsenic metabolism in the population of inner Mongolia, China, Toxicol. Sci. 155 (2017) 124–134.
[28] H. Yan, M. Li, L. Cao, H. Chen, H. Lai, Q. Guan, et al., A robust qualitative transcriptional signature for the correct pathological diagnosis of gastric cancer, J. Transl. Med. 17 (2019) 63.
[29] G. Hong, P. Zeng, N. Li, H. Cai, Y. Guo, X. Li, et al., A qualitative analysis based on relative expression orderings identifies transcriptional subgroups for Alzheimer's disease, Curr. Alzheimer Res. 16 (2019) 1175–1182.
[30] C. Cui, K. Chakraborty, X.A. Tang, G. Zhou, K.Q. Schoenfelt, K.M. Becker, et al., Neutrophil elastase selectively kills cancer cells and attenuates tumorigenesis, Cell 184 (2021) 3163–3177 e3121.
[31] H. Xian, Y. Li, B. Zou, Y. Chen, H. Yin, X. Li, et al., Identification of TIMELESS and RORA as key clock molecules of non-small cell lung cancer and the comprehensive analysis, BMC Cancer 22 (2022) 107.
[32] W. Chen, T. Song, F. Zou, Y. Xia, J. Xing, W. Yu, et al., Prognostic and immunological roles of IL18RAP in human cancers, Aging (Albany NY) 15 (2023).
[33] L.F. Deng, Identification of immune-related hub genes in thymoma: defects in CD247 and characteristics of paraneoplastic syndrome, Front. Genet. 13 (2022) 895587.
[34] M. Zhang, L. Zeng, Y. Peng, B. Fan, P. Chen, J. Liu, Immune-related genes LAMA2 and IL1R1 correlate with tumor sites and predict poor survival in pancreatic adenocarcinoma, Future Oncol. 17 (2021) 3061–3076.

[35] J.H. Esensten, Y.A. Helou, G. Chopra, A. Weiss, J.A. Bluestone, CD28 costimulation: from mechanism to therapy, Immunity 44 (2016) 973–988.
[36] H. Wang, Q. Sun, W. Zhao, L. Qi, Y. Gu, P. Li, et al., Individual-level analysis of differential expression of genes and pathways for personalized medicine, Bioinformatics 31 (2015) 62–68.
[37] C.A. Croft, F.A. Moore, P.A. Efron, P.S. Marker, A. Gabrielli, L.S. Westhoff, et al., Computer versus paper system for recognition and management of sepsis in surgical intensive care, J. Trauma Acute Care Surg. 76 (2014) 311–317, discussion 318-319.
[38] M. Zhou, H. Fang, M. Du, C. Li, R. Tang, H. Liu, et al., The modulation of regulatory T cells via HMGB1/PTEN/beta-Catenin Axis in LPS induced acute lung injury, Front. Immunol. 10 (2019) 1612.
[39] S. Liang, M. Xing, X. Chen, J. Peng, Z. Song, W. Zou, Predicting the prognosis in patients with sepsis by a pyroptosis-related gene signature, Front. Immunol. 13 (2022) 1110602.
[40] Q. Zhang, Y. Hu, P. Wei, L. Shi, L. Shi, J. Li, et al., Identification of hub genes for adult patients with sepsis via RNA sequencing, Sci. Rep. 12 (2022) 5128.
[41] Z. Sun, T. Zhang, C. Ning, D. Shen, W. Pei, R. Zhou, et al., Construction of sepsis diagnostic models and identification of macrophage subpopulations based on pyroptosis-related genes, Shock 60 (2023) 1–10.
[42] W. Gu, D. Wen, H. Lu, A. Zhang, H. Wang, J. Du, et al., MiR-608 exerts anti-inflammatory effects by targeting ELANE in monocytes, J. Clin. Immunol. 40 (2020) 147–157.
[43] R.A.C. Oliveira, D.O. Imparato, V.G.S. Fernandes, J.V.F. Cavalcante, R.D. Albanus, R.J.S. Dalmolin, Reverse engineering of the pediatric sepsis regulatory network and identification of master regulators, Biomedicines 9 (2021).
[44] J.A. García, H. Volt, C. Venegas, C. Doerrier, G. Escames, L.C. López, et al., Disruption of the NF-κB/NLRP3 connection by melatonin requires retinoid-related orphan receptor-α and blocks the septic response in mice, Faseb. J. 29 (2015) 3863–3875.
[45] E. Hams, J. Roberts, R. Bermingham, P.G. Fallon, Functions for retinoic acid-related orphan receptor alpha (RORα) in the activation of macrophages during lipopolysaccharide-induced septic shock, Front. Immunol. 12 (2021) 647329.
[46] C. He, Y. Zhang, Y. Duan, J. Yu, B. Luo, N. Jiang, et al., [Identification of differentially expressed genes and pathways changing in neutrophils of patients with sepsis by bioinformatics analysis], Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi 35 (2019) 481–490.
[47] Y. Jiang, Q. Miao, L. Hu, T. Zhou, Y. Hu, Y. Tian, FYN and CD247: key genes for septic shock based on bioinformatics and meta-analysis, Comb. Chem. High Throughput Screen. 25 (2022) 1722–1730.
[48] M. Chen, X. Chen, Y. Hu, X. Cai, Screening of key genes related to the prognosis of mouse sepsis, Biosci. Rep. 40 (2020).
[49] H. Ding, Y. Li, S. Chen, Y. Wen, S. Zhang, E. Luo, et al., Fisetin ameliorates cognitive impairment by activating mitophagy and suppressing neuroinflammation in rats with sepsis-associated encephalopathy, CNS Neurosci. Ther. 28 (2022) 247–258.
[50] M. Luo, H. Wang, K. Liu, M. Liu, S. Tan, Y. Zhu, et al., IL-1R1 blockade attenuates liver injury through inhibiting the recruitment of myeloid-derived suppressor cells in sepsis, Biochem. Biophys. Res. Commun. 620 (2022) 21–28.
[51] Y. Sun, J. Xie, J.C. Anyalebechi, C.W. Chen, H. Sun, M. Xue, et al., CD28 agonism improves survival in immunologically experienced septic mice via IL-10 released by Foxp3(+) regulatory T cells, J. Immunol. 205 (2020) 3358–3371.
[52] Y. Yao, J. Zhao, J. Hu, H. Song, S. Wang, Y. Wang, Identification of a four-gene signature for diagnosing paediatric sepsis, BioMed Res. Int. 2022 (2022) 5217885.
[53] F.C. Gong, R. Ji, Y.M. Wang, Z.T. Yang, Y. Chen, E.Q. Mao, et al., Identification of potential biomarkers and immune features of sepsis using bioinformatics analysis, Mediat. Inflamm. 2020 (2020) 3432587.
[54] H. Wang, J. Huang, W. Yi, J. Li, N. He, L. Kang, et al., Identification of immune-related key genes as potential diagnostic biomarkers of sepsis in children, J. Inflamm. Res. 15 (2022) 2441–2459.